

基于多时相遥感影像的水稻种植信息提取

张红华, 赵威成, 刘强凯 (黑龙江科技大学, 黑龙江哈尔滨 150022)

摘要 获取水稻种植信息对于指导水稻生产, 监测作物生长及合理分配水资源具有重要意义。针对基于单时相影像提取水稻信息精度有限, 以 Sentinel-2A/B 多时相影像为数据源, 构建 NDVI、EVI、NDWI 和光谱特征 4 种时序特征数据集并设计 6 种试验方案, 结合随机森林算法对水稻种植信息进行提取。结果表明, NDVI、EVI 时序曲线可以较好反映出水稻生育期的物候特征, 不同地类的光谱时序曲线和 NDWI 时序曲线可分离度较高, 有利于提高分类精度; 基于 NDVI 时序数据集的分类精度最低, 基于光谱时序数据集的分类精度最高, 总体精度达 95.559 0%, Kappa 系数为 0.943 3, 与基于 NDVI 的分类结果相比, 总体精度、Kappa 系数、水稻生产者精度和用户精度分别提高了 3.530 4%、0.044 9、8.64% 和 3.36%, 水稻与旱地的混发现象得到有效抑制。该研究为区域水稻种植信息精确提取在数据源选择、时序特征构建方面提供了一种新的思路和技术手段。

关键词 Sentinel-2A/B; 多时相; 时序特征; 水稻

中图分类号 S 127 文献标识码 A

文章编号 0517-6611(2022)07-0234-05

doi: 10.3969/j.issn.0517-6611.2022.07.056



开放科学(资源服务)标识码(OSID):

Rice Planting Information Extraction Based on Multi-temporal Remote Sensing Images

ZHANG Hong-hua, ZHAO Wei-cheng, LIU Qiang-kai (Heilongjiang University of Science and Technology, Harbin, Heilongjiang 150022)

Abstract Obtaining rice planting information is of great significance for guiding rice production, monitoring crop growth and rational allocation of water resources. In view of the limited accuracy of extracting rice information based on single temporal image, four time-series feature data sets of NDVI, EVI, NDWI and spectral features were created based on sentinel-2A/B multi-temporal images. Six experimental schemes were designed to extract rice planting information combined with random forest algorithm. The results showed that NDVI and EVI time series could better reflect the phenological characteristics of rice growth period, and the spectral time series and NDWI time series of different land types had a high degree of separation, which was conducive to improve the classification accuracy; the classification accuracy based on NDVI time series dataset was the lowest, and the classification accuracy based on spectral time series dataset was the highest, the overall accuracy was 95.559 0%, and the Kappa coefficient was 0.943 3. Compared with the classification results based on NDVI, the overall accuracy, Kappa coefficient, rice producer accuracy and user accuracy were improved by 3.530 4%, 0.044 9, 8.64% and 3.36%, respectively. And the mixing of rice and dry land was effectively controlled. This research provided a new idea and technical means for accurate extraction of regional rice planting information in data sources selection and time series feature construction.

Key words Sentinel-2A/B; Multi-temporal; Timing characteristics; Rice

准确获取区域水稻种植结构信息, 对于作物产量估计、种植结构调整具有重要意义^[1]。与传统的田间调查与统计汇总方法相比, 利用遥感技术提取作物种植结构, 更加直观和准确^[2]。水稻常与背景信息在空间上相互交错, 由水稻植株、水体和土壤的混合地物组成, 这种组合的季相变换带来的光谱信息差异可作为区分水稻田和其他地物的重要依据。作物生育期内, 仅基于单时相影像难以将水稻与其他作物区分, 可利用多时相影像获取作物的时间序列特征来提取水稻信息。以往研究多是通过构建归一化植被指数(Normalized Difference Vegetation Index, NDVI)时序数据集来进行作物的提取, 如利用 MODIS NDVI 时间序列实现作物的识别^[3-5], 基于 Sentinel-2^[6-8] 或 GF-1^[9-11] 数据构建 NDVI 时间序列, 基于多源影像构建 NDVI 时间序列等^[12]。然而, 目前利用多时相数据进行水稻提取多局限于构建 NDVI 时序数据, 对其他时序特征很少关注和研究。鉴于此, 笔者以黑龙江省五常市龙凤山镇为研究区, 利用作物整个生育期内的多时相 Sentinel-2A/B 遥感影像构建 4 种不同的时间序列特征, 包括光谱特征、归一化植被指数、增强

植被指数(Enhanced vegetation index, EVI)、归一化水体指数(Normalized difference water index, NDWI); 同时设计 6 种时序特征组合方案, 结合随机森林分类方法提取水稻空间分布信息, 以期提高区域尺度的水稻种植信息的提取精度, 为精准农情监测提供参考。

1 研究区及数据来源

该研究以黑龙江省五常市龙凤山镇为研究区(图 1), 五常市位于黑龙江省南部(44° 04' ~ 45° 26' N, 126° 33' ~ 128° 14' E), 是五常大米主产地。五常市属中温带大陆性季风气候, 夏短冬长, 寒暑悬殊, 年均气温 3~4 °C, 年均降水量 500~800 mm。由于气候特征的影响, 水稻主要分布于五常市西南的 C 型盆地中, 因其独特的地形条件, 有利于在 4—9 月份达到优质水稻所需要的活动积温。

根据五常水稻生育周期, 选取了涵盖 2021 年整个生长季(5—10 月)的 10 景 Sentinel-2 影像(表 1), 其中 4 景为 Sentinel-2A 数据, 6 景为 Sentinel-2B 数据。数据均来源于欧空局(ESA)数据中心(<https://scihub.copernicus.eu/>), 产品等级为 Level-2A 级, 其中红绿蓝和近红外 4 个波段空间分辨率为 10 m, 坐标系 WGS-84, 投影 UTM, Zone 52 N。该研究样本数据主要通过实地调查和依据天地图进行目视解译 2 种方式获取。于 2021 年 8 月对研究区进行样本采集, 包括地物类型、经纬度信息和生长状况。

基金项目 黑龙江省自然科学基金项目(JJ2017ZR0933); 黑龙江省省属高校 2019 年度基本科研业务费项目(Hkdqg201901)。

作者简介 张红华(1979—), 女, 辽宁朝阳人, 讲师, 在读博士, 从事地理信息获取与知识挖掘方面研究。

收稿日期 2021-11-15

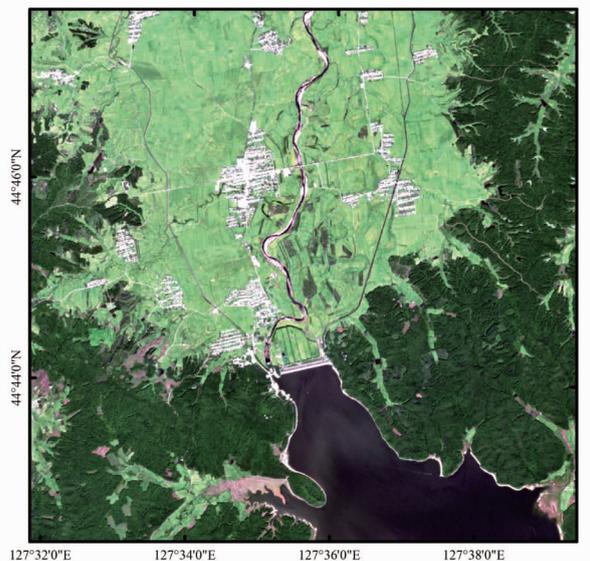


图 1 研究区影像

Fig. 1 Image of the study area

表 1 数据列表

Table 1 Data list

序号 Order number	获取日期 Sensing date	数据源 Data source	云量 Cloud amount/%
1	2021-05-10	Sentinel-2B	0
2	2021-05-20	Sentinel-2B	0
3	2021-06-19	Sentinel-2B	5
4	2021-07-24	Sentinel-2A	0
5	2021-08-08	Sentinel-2B	0
6	2021-08-18	Sentinel-2B	3
7	2021-09-02	Sentinel-2A	0
8	2021-09-17	Sentinel-2B	0
9	2021-10-12	Sentinel-2A	0
10	2021-10-22	Sentinel-2A	0

2 研究方法

以 10 景不同时期的 Sentinel-2A/B 影像为数据源,构建时间序列数据集,包括光谱特征、NDVI、EVI、NDWI,分析其时序曲线特征,利用随机森林算法对水稻种植信息进行提取并评定其精度。

2.1 时序特征数据集的构建

2.1.1 时序光谱数据集。该研究利用红绿蓝和近红外 4 个 10 m 波段构建时序光谱数据集,10 景影像共计 40 个波段,分析不同地类光谱曲线特征,进而完成分类。

2.1.2 NDVI 时序数据集。由于植被的红边效应,将红光波段与近红外波段进行组合形成的植被指数可以较好反应地表植被的生长状况,其中归一化植被指数(NDVI)是遥感影像分类中应用最广泛的植被指数,是植被生长状态和植被分布密度的最佳指数因子。NDVI 时间序列可以反映出作物从播种、生长、成熟到收割的作物生长过程,以及植被覆盖度的变化,并且消除一些辐射错误。NDVI 的计算方法如式(1)所示:

$$NDVI = (\rho_{nir} - \rho_{red}) / (\rho_{nir} + \rho_{red}) \quad (1)$$

式中, ρ_{nir} 为近红外波段的反射率; ρ_{red} 为红波段的反射率。

2.1.3 EVI 时序数据集。增强植被指数(EVI)通过加入蓝色波段以增强植被信号,矫正土壤背景和气溶胶散射的影响,该指数常用于植被茂密区,计算方法见式(2)。

$$EVI = 2.5 \left(\frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + 6\rho_{red} - 7.5\rho_{blue} + 1} \right) \quad (2)$$

式中, ρ_{nir} 为近红外波段的反射率; ρ_{red} 为红波段的反射率。 ρ_{blue} 为蓝波段的反射率。

2.1.4 NDWI 时序数据集。归一化水体指数(NDWI)是基于绿波段与近红外波段的归一化比值指数。该 NDWI 一般用来提取影像中的水体信息,效果较好。因水稻特有的水生特点,使得其生长环境离不开水体,所以引入 NDWI,以便对水稻和其他植被或作物进行区分,NDWI 的计算方法见式(3)。

$$NDWI = (\rho_{green} - \rho_{nir}) / (\rho_{green} + \rho_{nir}) \quad (3)$$

式中, ρ_{green} 为绿波段的反射率; ρ_{nir} 为近红外波段的反射率。

2.2 随机森林算法 随机森林是以 CART 决策树为基本分类器的集成学习模型^[13],采用自助抽样技术随机抽取大约 2/3 的样本作为训练样本子集,未被抽中的 1/3 称为袋外样本(Out-Of-Bag, OOB),OOB 数据可以通过内部交叉验证的方法估算每一棵决策树的泛化误差,将随机森林中所有决策树的 OOB 误差估计取平均值即可获得整个随机森林的泛化误差估计。当输入待分类的样本时,最终的分类结果由单个决策树的输出结果投票决定。随机森林克服了决策树过拟合问题,对噪声和异常值有较好的容忍性,在遥感影像信息提取方面取得了较好的分类效果^[14-16]。

2.3 试验方案 该研究设计 6 种试验方案进行对比研究(表 2),通过不同方案之间的对比,探索提高水稻信息提取精度的方法。

表 2 试验方案

Table 2 Experimental schemes

试验方案 Experimental scheme	时序特征组合 Sequence features combination	试验方案 Experimental scheme	时序特征组合 Sequence features combination
1	NDVI	4	NDVI+NDWI
2	EVI	5	EVI+NDWI
3	NDWI	6	光谱特征

3 结果与分析

3.1 时序特征曲线 利用 10 期 5—10 月的遥感影像,分地类计算每期训练样本光谱特征、DNVI、EVI、NDWI 的平均值,绘制时序曲线,如图 2~5 所示。

从图 2 可以看出,水体和建设用地的时序光谱曲线平稳,不同时期光谱反射率变化幅度较小,与绿色植被(林地、水稻、旱地)可分离度较高。水稻、旱地和林地 在 7—8 月反射率值较高,整个生长季波谱变化幅度较大,但从时序曲线上来看,绿色植被间也表现了明显的差异性,有利于提高其分类精度。从图 3 的 NDVI 时序曲线来看,5 种地类均有独特的时序特征,林地、建设用地和水体可分离性较高,水稻和旱地在 7—8 月间 NDVI 值极为接近,可分离性稍差,该时期

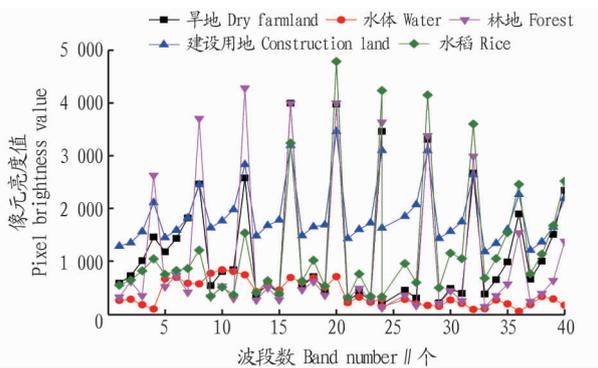


图2 时序光谱曲线

Fig.2 Sequence spectral curves

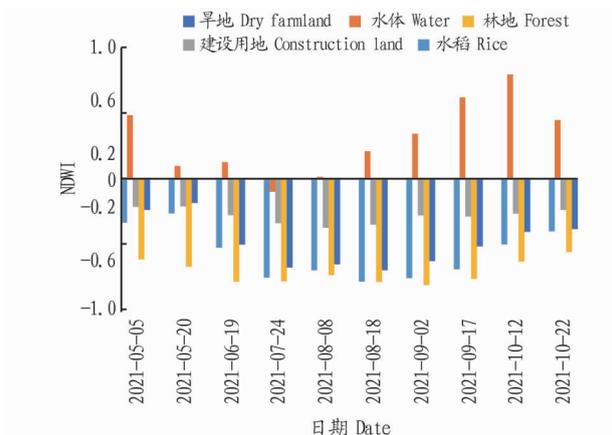


图5 NDWI 时序曲线

Fig.5 NDWI sequence curves

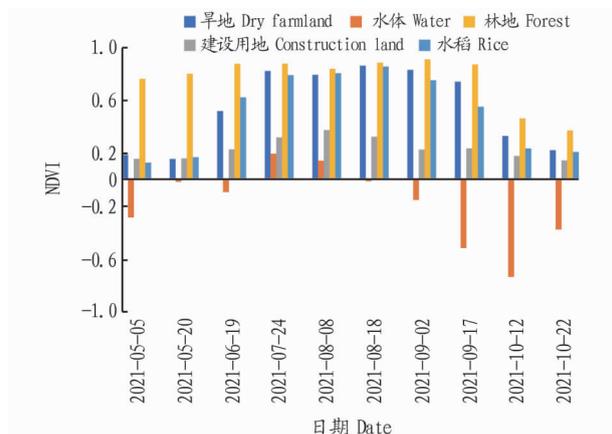


图3 NDVI 时序曲线

Fig.3 NDVI sequence curves

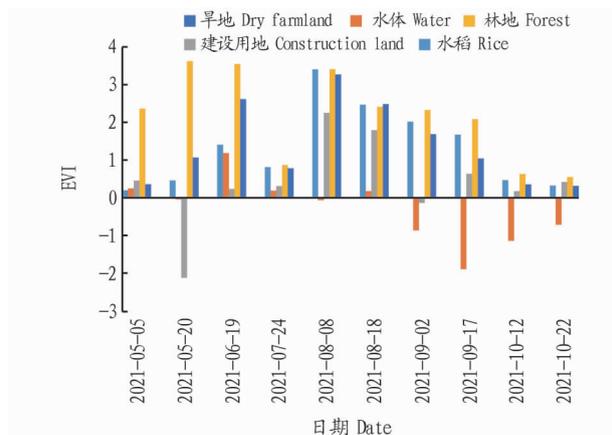


图4 EVI 时序曲线

Fig.4 EVI sequence curves

为水稻拔节期和抽穗期,生长旺盛,ENVI 值较高,最大值达 0.848 0。由图 4 可知,各地类 EVI 时序曲线差异较明显,但在 7—8 月间,水稻、旱地和林地的 EVI 值差异非常小,在 7 月 24 日、8 月 8 日和 8 月 18 日的遥感影像上,绿色植被可分离性较差。由图 5 中 NDWI 时序曲线可知,各地类间特征差异显著,在生长季易混分的水稻和旱地也具有较好的可分离性。总体来看,4 类时序曲线仅在少数时相上存在水稻和旱地的分离度较低的情况,但各地类均有独特的时序特征,有利于分类精度的提高。

3.2 随机森林分类结果及精度评价 依据 6 种特征组合方案进行随机森林分类,提取结果如图 6 所示。该研究将研究区分为水稻、旱地、林地、建设用地和水体 5 种类型。通过与原始图像对比可以看出,不同方案中旱地的提取结果存在差异,主要原因是水稻和旱地的混分;水体中大面积水域(即龙凤山水库)提取精度高,而河流和沟渠提取精度较低,存在建设用地和水体混分的现象;林地提取结果较理想。

利用验证样本计算混淆矩阵,通过总体精度、Kappa 系数、各类别的生产者精度(Producer's Accuracy, PA)、用户精度(User's Accuracy, UA)来评价不同试验方案的分类精度(表 3)。从总体精度和 Kappa 系数来看,方案 1 和 2 的分类精度相近,方案 3、4 和 5 的分类精度接近并高于方案 1、2,方案 6 总体精度最高,与总体精度最低的方案 1 相比,总体精度和 Kappa 系数分别提高了 3.530 4%和 0.044 9,说明利用时序光谱特征进行分类的总体精度最高。由各地类的生产者精度和用户精度可知,方案 1 即基于时序 NDVI 分类,旱地与水稻混分比例较高,旱地与水稻的生产者精度均低于 90.00%,旱地的用户精度低于 70.00%,在方案 2 中混分现象没有明显改善,方案 3 中水稻和旱地的分类精度有较大提升,水稻的 PA 和 UA 分别提升了 4.29%和 2.68%,旱地分别提升了 4.26%和 8.95%;方案 4 和 5 的结果与方案 3 相比,分类精度提高有限,甚至方案 4 整体分类精度略低于方案 3,说明在 NDWI 的基础上加入 NDVI 或 EVI,分类精度并没有明显提高;方案 6 总体精度最高,水稻和旱地的混分现象得到有效抑制,与方案 1 相比,水稻的 PA 和 UA 分别提升了 8.64%和 3.36%。综上,方案 6 即利用时序光谱特征进行分类,各地类的分类精度及总体精度均较高,方案 3 即基于时序 NDWI 的分类结果也较为理想。

4 结论

利用覆盖水稻完整生育期的多时相 Sentinel-2A/B 数据,构建 NDVI、EVI、NDWI 和光谱特征时间序列数据集并设计 6 种试验方案,结合随机森林算法进行五常市龙凤山镇水稻信息提取,同时计算混淆矩阵来评价不同实验方案的水稻种植信息提取精度,主要结论如下:

(1) NDVI、EVI 时序曲线可以明显反映出水稻生长季的

物候特征;各类地物的光谱时序曲线和 NDWI 时序曲线均具有的特征,有利于提高分类精度。

地混分比例较高,分类精度较低;基于光谱特征时序数据集的分类结果,总体精度、Kappa 系数、水稻生产者精度和用户精度均较高,同时水稻与旱地的混分现象得到有效控制。

(2) 基于 NDVI、EVI 时序数据集的分类结果,水稻与旱

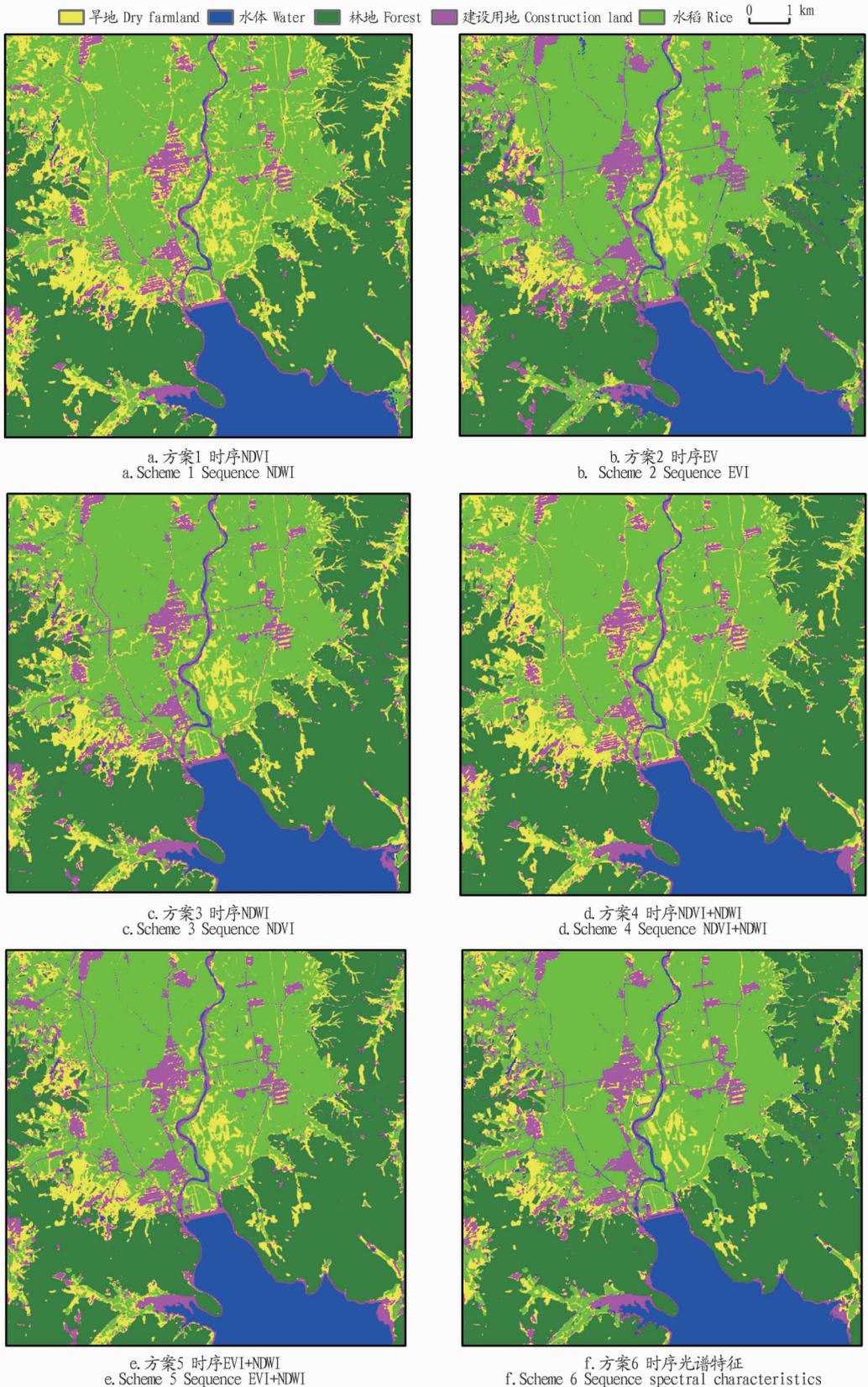


图 6 不同试验方案分类结果

Fig. 6 Classification results of different experimental schemes

表 3 分类精度

Table 3 Classification accuracy

%

类别 Category	方案 1 Scheme 1		方案 2 Scheme 2		方案 3 Scheme 3		方案 4 Scheme 4		方案 5 Scheme 5		方案 6 Scheme 6	
	PA	UA										
旱地 Dry farmland	85.44	68.74	79.51	74.58	89.70	77.69	91.49	73.08	88.47	76.33	93.39	79.81
水体 Water	89.04	97.13	90.47	97.49	90.02	98.28	90.17	98.36	90.09	98.93	92.12	97.30
建设用地 Construction land	96.53	92.56	97.81	89.04	96.90	90.15	97.63	94.19	98.72	91.39	96.26	98.05
林地 Forest	96.69	99.03	97.87	98.62	97.07	98.75	97.02	98.94	97.16	98.94	96.50	98.98
水稻 Rice	89.13	94.87	90.60	93.62	93.42	97.55	91.48	97.99	93.42	97.61	97.77	98.23
总体精度 Overall accuracy//%	92.028 6		92.448 9		93.933 9		93.821 8		94.102 0		95.559 0	
Kappa 系数 Kappa coefficient	0.898 4		0.903 6		0.922 6		0.921 3		0.924 8		0.943 3	

由于该研究采用的是基于像素的分类方法,分类结果存在一定的“椒盐现象”,未来应将时序数据集与基于对象的分类方法结合,减少“同物异谱”和“同谱异物”的产生。

参考文献

- [1] 陈安旭,李月臣. 基于 Sentinel-2 影像的西南山区不同生长期水稻识别[J]. 农业工程学报,2020,36(7):192-199.
- [2] 麦丽素,乌兰吐雅. 基于 Sentinel-2 数据的水稻面积提取方法比较分析[J]. 北方农业学报,2019,47(5):119-126.
- [3] 周思,何祺胜,刘宝柱. 基于 MODIS 的黑龙省农作物种植结构提取研究[J]. 地理空间信息,2018,16(1):79-82.
- [4] 郭昱杉,刘庆生,刘高焕,等. 基于 MODIS 时序 NDVI 主要农作物种植信息提取研究[J]. 自然资源学报,2017,32(10):1808-1818.
- [5] 朱彤,张学霞,王士远,等. 基于物候特征和混合光谱信息的春玉米种植面积提取[J]. 沈阳农业大学学报,2017,48(3):328-337.
- [6] 王利军,郭燕,贺佳,等. 基于决策树和 SVM 的 Sentinel-2A 影像作物提取方法[J]. 农业机械学报,2018,49(9):146-153.
- [7] 毕恺艺,牛铮,黄妮,等. 基于 Sentinel-2A 时序数据和面向对象决策树方法的植被识别[J]. 地理与地理信息科学,2017,33(5):16-20.
- [8] 杜保佳,张晶,王宗明,等. 应用 Sentinel-2A NDVI 时间序列和面向对

象决策树方法的农作物分类[J]. 地球信息科学学报,2019,21(5):740-751.

- [9] 张海东,田婷,张青,等. 基于 GF-1 影像的耕地地块破碎区水稻遥感提取[J]. 遥感技术与应用,2019,34(4):785-792.
- [10] 杨闫君,占玉林,田庆久,等. 基于 GF-1/WFV NDVI 时间序列数据的作物分类[J]. 农业工程学报,2015,31(24):155-161.
- [11] 万丛,梁治华,张锦水. 基于高分一号宽幅时序影像的冬小麦分布识别研究[J]. 安徽农业科学,2020,48(23):256-259.
- [12] 王碧晴,韩文泉,许驰. 基于图像分割和 NDVI 时间序列曲线分类模型的冬小麦种植区域识别与提取[J]. 国土资源遥感,2020,32(2):219-225.
- [13] 张磊,官兆宁,王启为,等. Sentinel-2 影像多特征优选的黄河三角洲湿地信息提取[J]. 遥感学报,2019,23(2):313-326.
- [14] 何云,黄肿,李贺,等. 基于 Sentinel-2A 影像特征优选的随机森林土地覆盖分类[J]. 资源科学,2019,41(5):992-1001.
- [15] 冯文卿,眭海刚,涂煜辉,等. 高分辨率遥感影像的随机森林变化检测方法[J]. 测绘学报,2017,46(11):1880-1890.
- [16] 詹国旗,杨国东,王凤艳,等. 基于特征空间优化的随机森林算法在 GF-2 影像湿地分类中的研究[J]. 地球信息科学学报,2018,20(10):1520-1528.

(上接第 226 页)

资源、公共服务、政策支持,二者互为需求和反馈。当城镇综合系统需要产业系统作出调整,扩大木业下游企业规模时,产业系统作出相应反馈,延伸产业链条;当产业系统需要提升人力资源素质、加大财政支持时,城镇系统也作出相应反馈。系统内部的主动适应与调整,可以大大提升发展效率,延缓熵增的过程。

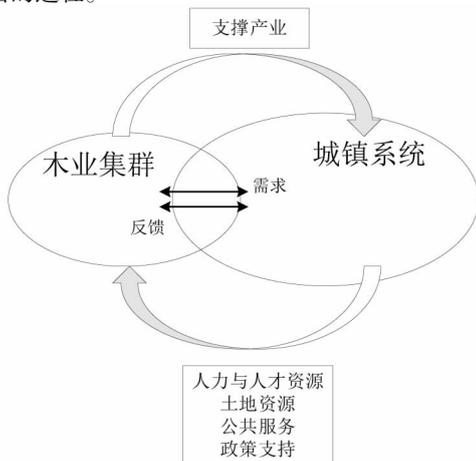


图 4 区域经济空间组织动态反馈与调整系统

Fig. 4 Dynamic feedback and adjustment system of regional economic spatial organization

参考文献

- [1] 马歇尔. 经济学原理[M]. 陈良壁,译. 北京:商务印书馆,2005.
- [2] 王缉慈. 创新的空间:企业集群与区域发展[M]. 北京:北京大学出版社,2001.
- [3] 陈晖涛. 福建省农村城镇化模式选择研究[D]. 福州:福建农林大学,2014.
- [4] 林珊,林发彬. 新型城镇化进程与产业集聚支撑能力研究:以福建省为例[J]. 亚太经济,2018(4):94-102.
- [5] 张治栋,孟东涛. 长江经济带产业集聚推动城镇化了吗?——基于 108 个地级市 2005-2015 年数据的实证分析[J]. 华东经济管理,2018,32(6):72-79.
- [6] 伍骏骞,何伟,储德平,等. 产业集聚与多维城镇化异质性[J]. 中国人口·资源与环境,2018,28(5):105-114.
- [7] 马国勇,蔡玲松. 城镇化与产业集聚交互作用机理研究[J]. 哈尔滨工业大学学报(社会科学版),2019,21(5):127-134.
- [8] 纪玉俊,郝婷婷. 民族地区制造业集聚与城镇化的互动效应[J]. 中南民族大学学报(人文社会科学版),2021,41(3):71-80.
- [9] 宋瑛,廖蔓,王亚飞. 制造业集聚对新型城镇化的影响研究:基于空间溢出效应的视角[J]. 重庆大学学报(社会科学版),2019,25(6):1-13.
- [10] 朱喜安,张秀,李浩. 中国高新技术产业集聚与城镇化发展[J]. 数量经济技术经济研究,2021,38(3):84-102.
- [11] 陈斌. 长三角城镇化、产业集聚与区域创新承载力的耦合关系[J]. 南通大学学报(社会科学版),2020,36(1):42-49.
- [12] 张贵先. 重庆市产业集聚与城镇化互动发展模式研究[D]. 重庆:西南大学,2012.
- [13] 曹炳汝,孙巧. 产业集聚与城镇空间格局的耦合关系及时空演化:以长三角区域为例[J]. 地理研究,2019,38(12):3055-3070.
- [14] 王琦. 产业集聚与区域经济空间耦合机理研究[D]. 长春:东北师范大学,2008.
- [15] 闫芳,汤振兴. 基于产业集聚的河南省新型城镇化发展水平评价[J]. 中国农业资源与区划,2019,40(10):188-194.
- [16] 谭跃进,邓宏钟. 复杂适应系统理论及其应用研究[J]. 系统工程,2001,19(5):1-6.